

# Conditional simulations of the extremal $t$ process: application to fields of extreme precipitation

Aurélien Bechler<sup>a,b,c</sup>, Liliane Bel<sup>a,b</sup>, Mathieu Vrac<sup>c</sup>

<sup>a</sup>*INRA, UMR 518 Math. Info. Appli., F-75005 Paris, France*

<sup>b</sup>*AgroParisTech, UMR 518 Math. Info. Appli., F-75005 Paris, France.*

<sup>c</sup>*LSCE-IPSL, Centre d'Etudes de Saclay, Orme des Merisiers, Bat. 701 91191 Gif-sur-Yvette, France*

---

## Abstract

The last decade has seen max-stable processes emerge as a powerful tool for the statistical modeling of spatial extremes and there are increasing works using them in a climate framework. One recent utilization of max-stable processes in this context is for conditional simulations that provide empirical distribution of a spatial field conditioned by observed values in some locations. In this work conditional simulations are investigated for the extremal  $t$  process taking benefits of its spectral construction.

The methodology of conditional simulations proposed by Dombry et al. (2013) for Brown-Resnick and Schlather models is adapted for the extremal  $t$  process with some original improvements which enlarge the possible number of conditional points. A simulation study enables to highlight the role of the different parameters of the model and to emphasize the importance of the steps of the algorithm.

An application is performed on precipitation data in the south of France where extreme precipitation events (Cevenol) may generate major floods. This shows that the model and the algorithm perform well provided the stationary assumptions are fulfilled.

### *Keywords:*

Max-stable processes, extremal  $t$  processes, spatial extremes, conditional simulation, extreme precipitation.

---

## 1. Introduction

Extreme events are one of the major demonstrations of complex environmental systems. As they are inherently rare and often unexpected, society is not well prepared to face them. One important feature of extreme events is that their probability decreases rapidly with increasing intensity, while damage increases strongly as does the cost of protection against such events. Both analyses of observational data and projections from climate models suggest that the occurrence and magnitudes of some extreme events will increase in the future (IPCC (2011)). A potential consequence is an increasing number of catastrophes such as flooding, storms or crop failures. In addition to their direct effects and costs in terms of human lives, building destruction or food security, these events will also have indirect impacts such as increasing costs for building

or strengthening infrastructures, in addition to higher insurance premiums, and before even mentioning the consequences of global population displacements (IPCC (2011)). A better understanding and modeling of extreme events is therefore a crucial need in many economic aspects and societal contexts for both public and private decision makers.

From a mathematical point of view, the suitable probabilistic framework for modeling the frequency and intensity of those special events is the statistical theory of extreme values. The extreme value theory is well established for univariate stationary series. There is now a growing effort to develop methods and models to deal with complex data involved in applications where space-time problems and non-stationarity are the rule rather than the exception.

Multivariate extreme distributions have been used in climate and environmental applications in order to take into account extremal dependence (Coles and Tawn (1994), De Haan and De Ronde (1998), Schlather and Tawn (2003), Fawcett and Walshaw (2012)). For the study of extremes in a spatial context, max-stable processes are a relevant framework for modeling and inference. Introduced by De Haan (1984) different parametric models have been later proposed by Smith (1990), Schlather (2002) or Kabluchko et al. (2009). The extremal  $t$  process (Padoan (2011), Opitz (2013)) is a model of that family that extends the Schlather model. It has a supplementary parameter that lets the extremal coefficient function vary in a wider range. Max-stable processes have been used in applications for modeling rainfall (Buishand et al. (2008), Davison et al. (2012)), temperature (Davison and Gholamrezaee (2012), Fuentes et al. (2013)), snowfall (Gaume et al. (2013)) or snow depth (Blanchet and Davison (2011)). A main objective of these works is to provide return periods of rare events for prevention purposes even in locations where no observations are available.

The aim of the present paper is to further investigate conditional simulations of extreme precipitation. Conditional simulations provide empirical distributions of a random field knowing the values observed at some given locations. Consequently, they allow us to statistically generate extreme values at locations without any recording, in agreement with surrounding observations. This capacity to generate time series at observation-free locations is of major importance for some impact models, such as hydrological models that require input data, e.g. precipitation values, correctly spatially distributed. The spatial coverage of the meteorological inputs of such models has to be as precise as possible. Furthermore, when other techniques provide individual values or indicators, conditional simulations provide distributions at every new locations which allows us to get numerous statistical quantities as e.g. quantile, dispersion indicators, etc... The conditional simulation techniques are then of major importance.

For Gaussian and Boolean fields, algorithms for conditional simulations are well established (Lantu  joul (2002)). They have been used for representing physical phenomena by integrating heterogeneous data and assigning confidential limits to the estimates. For example, they have been used for predicting the geometry of oil reservoirs or the grade of mining blocks (Chil  s and Delfiner (1999)) and are used more and more in the environmental sciences (Mariethoz et al. (2009)). Concerning max-stable

processes, Wang and Stoev (2011) consider max-linear random fields, a special case of max-stable processes, and provide an exact expression of the conditional distributions as well as an efficient algorithm for the simulation. In a further work, Dombry and Eyi-Minko (2013) derive the regular conditional distribution for the very general class of max-infinitely divisible processes which encompasses the max-stable processes. This result has been applied then to provide a methodology and an algorithm for Brown-Resnick and Schlather processes (Dombry et al. (2013)). However, for processes which are not regular, that is the exponent measure is not absolutely continuous with respect to Lebesgue measure, the conditional distribution is not explicit. Oesting and Schlather (2014) discuss the frameworks where Wang and Stoev (2011) and Dombry et al. (2013) can be applied, and, in case of non regular models, propose an algorithm based on martingale arguments.

In this paper we adapt the methodology of Dombry et al. (2013) to the extremal  $t$  process. In this case all the calculations of the algorithm can be derived explicitly, nevertheless some steps still present computational issues. In particular, one of the steps of the algorithm involves an acceptance/rejection procedure for which the rate of rejection increases very fast with the number of conditioning points. This limiting point makes the algorithm inapplicable when the number of conditioning points is greater than approximately 30. To bypass the issue, we propose some original modifications of the algorithm to replace the acceptance/rejection step by a Gibbs sampler which speeds up the calculations and enables a large number of conditioning points. Despite this improvement, some difficulties persist in the algorithm, mostly due to the calculation of high dimensional integrals that have to be approximated. Then to get a better understanding of the algorithm mechanisms, a simulation study is led to assess the influence of the parameters from the extremal  $t$  process and to evaluate the importance of the most problematic and time-consuming step of the algorithm.

We focus on the French Mediterranean basin, which is a key spot of occurrences of meteorological extremes such as heavy precipitation (e.g. Ducrocq et al. (2008); Lebeaupin et al. (2006); Beaulant et al. (2011); Vrac and Yiou (2010) or Vrac et al. (2012)). Indeed, major extreme precipitation events are regularly observed in this region near the Cevennes mountains. The modeling and the forecast of this extreme precipitation – the so-called “Cevenol events” – are of major importance for hydrological studies in this complex terrain since they often trigger major floods in this region (e.g. Ducrocq et al. (2008)).

The structure of this article is the following. In Section 2, a short reminder on max-stable processes is provided and the extremal  $t$  process is described. Section 3 contains theoretical considerations on conditional simulations of an extremal  $t$  process, as well as the associated algorithm and its improvement. An extensive simulation study is carried out in Section 5 to evaluate the performances of the algorithm. The use of the continuous ranking probability score and the quantile score (Gneiting and Raftery (2007)) enables us to assess the accuracy of our methodology and to compare it to other models. This evaluation allows us to quantify the importance of the model parameters

and of the algorithm steps. Section 6 displays an application to maximum precipitation data in the south of France where extreme precipitation may cause major floods. Some comments and conclusions are given in section 7.

## 2. Extreme Value Theory

### 2.1. Max-Stable Process

We will note vectors in plain text and vectors of vectors in bold. Let us consider a stationary random field  $\mathbf{Z} = \{Z(x)\}_{x \in \mathbb{R}^d}$  with a given distribution  $\mathcal{H}_0$  and  $(Z_1, \dots, Z_n)$  independent copies of  $Z$ . If there exist sequences  $a_n(x) \in \mathbb{R}$  and  $b_n(x) \in \mathbb{R}^+$  such that for all  $n \in \mathbb{N}$ ,

$$\left\{ \frac{\max_{i=1, \dots, n} Z_i(x) - a_n(x)}{b_n(x)} \right\}_{x \in \mathbb{R}^d} \quad (1)$$

follows the distribution  $\mathcal{H}_0$ , then  $\mathbf{Z} = \{Z(x)\}_{x \in \mathbb{R}^d}$  is a max-stable process (De Haan (1984)).

As a consequence, all finite dimensional marginal distributions are max-stable, and in particular, the univariate margins  $Z(x)$  distribution belong to the Generalized Extreme Value (GEV) family:

$$\mathbb{P}(Z(x) \leq z) = \begin{cases} \exp\left(-\left[1 + \xi(x)\left(\frac{z - \mu(x)}{\sigma(x)}\right)\right]^{-1/\xi(x)}\right) & \text{if } \xi(x) \neq 0, \\ \exp\left(-\exp\left(\frac{z - \mu(x)}{\sigma(x)}\right)\right) & \text{if } \xi(x) = 0, \end{cases} \quad (2)$$

where  $\mu(x)$ ,  $\sigma(x)$  and  $\xi(x)$  are respectively the location, scale, and shape parameters at site  $x$ . Furthermore, the sign of  $\xi(x)$  will determine the family distribution  $Z(x)$  belongs to among Fréchet ( $\xi > 0$ ), Gumbel ( $\xi = 0$ ) or Weibull ( $\xi < 0$ ), and consequently the distribution tail characteristics.

Different max-stable models have been proposed and used in the past by several authors, e.g. Brown and Resnick (1977), Smith (1990), Schlather (2002) or Kabluchko et al. (2009).

Their margins are Fréchet distributed, i.e.

$$\mathbb{P}(Z(x) \leq z) = e^{-z^{-1}}. \quad (3)$$

### 2.2. The extremal $t$ process: definition and properties

The extremal  $t$  process (Padoan (2011)) belongs to the max-stable process family. Opitz (2013) proposed a spectral construction of the extremal  $t$  process which leads to the general expression

$$\mathbf{Z} = \{Z(x)\} = \left\{ m_\alpha^{-1/\alpha} \max_{i \geq 1} \xi_i W_i(x) \right\}, \quad \alpha > 0, \quad x \in \mathcal{X} \subset \mathbb{R}^d, \quad \text{with } \mathcal{X} \text{ a Borel set}, \quad (4)$$

with  $m_\alpha = \sqrt{\pi}^{-1} 2^{(\alpha-2)/2} \Gamma((\alpha+1)/2)$ ,  $\xi_i$  a Poisson point process on  $(0, +\infty)$  with intensity measure  $d\Lambda(t) = \alpha t^{-(\alpha+1)} dt$  and  $W_i$  independent copies of a stationary standard Gaussian process on  $\mathbb{R}^d$  with correlation function  $\rho(\cdot)$ .  $\mathbf{Z}$  margins are  $\alpha$ -Fréchet distributed, i.e.

$$\mathbb{P}(Z(x) \leq z) = e^{-z^{-\alpha}}. \quad (5)$$

Unlike the usual representation of max-stable processes with unit-Fréchet margins, we choose the  $\alpha$ -Fréchet representation (as in Opitz (2013)) to highlight the major role of the  $\alpha$ -parameter.

It has to be noted that for  $\alpha = 1$ , this representation leads to a Schlather process and hence the extremal  $t$  process can be considered as a generalization of the Schlather process, with a greater flexibility.

The parameter  $\alpha$  and the correlation function  $\rho(\cdot)$  determine the dependence structure. Furthermore, as will be seen in the next section,  $\alpha$  characterizes the spatial asymptotic dependence, that is the dependence in extreme values.

### 2.3. Extremal coefficient

For a max-stable process such that all margins are equally distributed the extremal coefficient function  $\theta(\cdot)$  (Schlather and Tawn (2003)) is defined as

$$\mathbb{P}(Z(x) < u, Z(x+h) < u) = [\mathbb{P}(Z(x) < u)]^{\theta(h)}. \quad (6)$$

$\theta(h) \in [1, 2]$  measures the strength of asymptotic dependence,  $\theta(h) = 1$  corresponds to the complete dependence case,  $\theta(h) = 2$  corresponds to the independence case. The extremal coefficient of the extremal  $t$  process depends on  $\alpha$  and is given by (Opitz (2013))

$$\theta(h) = 2t_{\alpha+1}(1 \mid \rho(h), (\alpha+1)^{-1}(1 - \rho(h)^2)), \quad (7)$$

with  $t_\nu(\cdot \mid \mu, \Sigma)$  the cdf (cumulative distribution function) of a Student distribution with the parameters  $\mu$ ,  $\Sigma$  and  $\nu$  (respectively location, shape and degree of freedom parameters).

In particular for the extremal  $t$  process with  $\alpha \geq 0$ ,  $\theta_\infty$ , the limit of  $\theta(h)$  when  $h$  tends to  $\infty$ , belongs to the interval  $]1.5, 2[$ , while  $\theta_\infty = 1 + \sqrt{2}/2$  ( $\approx 1.7$ ) for the Schlather process.

Consequently, thanks to this supplementary parameter  $\alpha$ , the extremal  $t$  process enables us to model a greater variety of situations for asymptotic dependence. Furthermore, this flexibility of the asymptotic dependency behavior is different from the one given by the class of general Brown-Resnick processes (Kablichko et al. (2009)).

## 3. Conditional simulations

Conditional simulations provide simulations of random fields at locations without observation, given observations in fixed locations, according to a model. The simulated fields are required to fit the observed values at the fixed locations. In contrast to simple simulations for which simulated values follow a target distribution – i.e.  $\mathbf{Z}$  is a

realization from a distribution  $\mathcal{L}$  –, conditional simulations take into account constraints expressed by the observations, i.e.

$$\begin{cases} \mathbf{Z} \text{ is a realization from a distribution } \mathcal{L}, \\ Z(x_i) = z_i, \forall z_i \text{ observations at the locations } x_i, i = 1, \dots, k. \end{cases}$$

Conditional simulations are powerful tools which provide not only individual values at new locations, but probabilistic distributions at those locations, given a chosen model. Those distributions allow the calculation of many statistical quantities as for example quantiles, dispersion indicators, etc. . . .

For some models, such as Gaussian or Boolean, algorithms for performing conditional simulations are quite straightforward and based on kriging techniques (Lantuéjoul (2002)). Dombry et al. (2013) proposed a general framework for conditional simulations of max-stable processes. The max-stable process  $\mathbf{Z}$  is written as the point-wise maximum over an infinity of continuous functions

$$Z(x) = \max_{i \geq 1} \phi_i(x). \quad (8)$$

Let  $\mathbf{x} = (x_1, x_2, \dots, x_k) \in \mathcal{X}^k$  the  $k$  conditioning points with  $\mathcal{X} \subset \mathbb{R}^d$ ,  $z = (z_1, z_2, \dots, z_k) \in \mathbb{R}^k$  such that  $Z(x_i) = z_i$  and  $\mathbf{s} = (s_1, s_2, \dots, s_m) \in \mathcal{X}^m$  the  $m$  locations the process  $\mathbf{Z}$  is to be simulated. For each conditioning point  $x_i$ , an unique function  $\phi_j$  almost surely exists, called extremal function and denoted  $\phi_j^+$  such that  $\phi_j^+(x_i) = z_i$ . One extremal function may be extremal for several conditioning points. Let  $\tau = (\tau_1, \dots, \tau_l)$  be a partition of the set  $\{x_1, \dots, x_k\}$  with  $l = |\tau|$  ( $\leq k$ ) such that

$$\phi_j^+(x_i) = z_i \text{ if } x_i \in \tau_j \text{ and } \phi_j^+(x_i) < z_i \text{ if } x_i \notin \tau_j, \quad (i = 1, \dots, k; j = 1, \dots, l). \quad (9)$$

The partition  $\tau$  is called a hitting scenario (Wang and Stoev (2011)). The conditional simulation algorithm is then divided into three steps. The first one consists in drawing a partition of  $\{x_1, \dots, x_k\}$  such that each class of the partition gathers conditioning points that have the same extremal function. Then, for each class in the drawn partition, each extremal function  $\phi_j^+$  is drawn according to the constraints (9) given by the conditional observations. Finally a number  $N$  sufficiently large of non extremal functions denoted  $\phi_j^-$  are drawn such that for all conditioning points they verify

$$\phi_j^-(x_i) < z_i \text{ for all } x_i \quad (i = 1, \dots, k; j = 1, \dots, N). \quad (10)$$

The simulated values on the set  $\mathbf{s}$  are obtained by taking the point-wise maximum of  $\{\phi_j^-(\mathbf{s}), \phi_j^+(\mathbf{s})\}$ . To perform the three steps of the algorithm, the conditional distribution of the partition, the  $\{\phi^+\}$  and the  $\{\phi^-\}$  are needed. Dombry et al. (2013) derived the expression of the conditional distributions for the Schlather and the Brown-Resnick process. The conditional distribution for the extremal  $t$  process is obtained thanks to a slight adaptation.

### 3.1. Conditional distribution for the extremal $t$ process

Let  $\mathbf{Z}$  be an extremal  $t$  process. It can be written as

$$Z(x) = \max_{i \geq 1} \phi_i^{(\alpha)}(x) \quad (11)$$

with

$$\phi_i^{(\alpha)}(x) = m_\alpha^{-1/\alpha} \xi_i W_i(x), \quad i = 1, 2, \dots \quad (12)$$

$\Phi^{(\alpha)} = \{\phi_i^{(\alpha)}\}_{i \geq 1}$  is a Poisson point process on  $\mathcal{C}_0$ , the space of positive continuous function on  $\mathcal{X} \subset \mathbb{R}^d$ . The intensity measure  $\Lambda_{\mathbf{x}}^{(\alpha)}$  of the point process  $\Phi^{(\alpha)}(\mathbf{x}) = \{\phi_i^{(\alpha)}(\mathbf{x})\}_{i \geq 1}$  on  $(0, +\infty)^k$  is

$$\Lambda_{\mathbf{x}}^{(\alpha)}(A) = \int_0^{+\infty} \mathbb{P}[\Phi^{(\alpha)}(\mathbf{x}) \in A] \xi^{-(\alpha+1)} d\xi, \quad A \subset (0, +\infty)^k \text{ a Borel set.} \quad (13)$$

The extremal  $t$  process fulfills the conditions necessary for  $\Lambda_{\mathbf{x}}^{(\alpha)}$  being absolutely continuous with respect to the Lebesgue measure (Dombry and Eyi-Minko (2013)). The density (Radon-Nikodym derivative)  $\lambda_{\mathbf{x}}^{(\alpha)}(z)$  such that  $\Lambda_{\mathbf{x}}^{(\alpha)}(dz) = \lambda_{\mathbf{x}}^{(\alpha)}(z) dz$  and the conditional density

$$\lambda_{\mathbf{s}|\mathbf{x},z}^{(\alpha)}(u) = \frac{\lambda_{\mathbf{s},\mathbf{x}}^{(\alpha)}(u, z)}{\lambda_{\mathbf{x}}^{(\alpha)}(z)}, \quad (\mathbf{s}, \mathbf{x}) \in \mathcal{X}^{m+k}, z \in (0, +\infty)^k, u \in \mathbb{R}^m. \quad (14)$$

are given by :

$$\lambda_{\mathbf{x}}^{(\alpha)}(z) = \frac{\pi^{\frac{\alpha-1}{2}}}{w_{k+\alpha-2}} |\Sigma_{\mathbf{x}}|^{-1/2} [z^T \Sigma_{\mathbf{x}}^{-1} z]^{-(\frac{k+\alpha}{2})} \Gamma((\alpha+1)/2)^{-1} \quad (15)$$

and

$$\lambda_{\mathbf{s}|\mathbf{x},z}^{(\alpha)}(u) = \frac{w_{k+\alpha-2}}{w_{k+m+\alpha-2}} \sqrt{|\Sigma_{\mathbf{s},\mathbf{x}}|^{-1} |\Sigma_{\mathbf{x}}|} \frac{[z^T \Sigma_{\mathbf{x}}^{-1} z]^{(k+\alpha)/2}}{[(u, z)^T \Sigma_{\mathbf{s},\mathbf{x}}^{-1} (u, z)]^{(k+m+\alpha)/2}}, \quad (16)$$

with  $w_d = \pi^{d/2} / \Gamma(d/2 + 1)$ ,  $|\cdot|$  the determinant operator and  $\Sigma_{\mathbf{x}}, \Sigma_{\mathbf{s},\mathbf{x}} = \begin{pmatrix} \Sigma_{\mathbf{s}} & \Sigma_{\mathbf{s}:\mathbf{x}} \\ \Sigma_{\mathbf{x}:\mathbf{s}} & \Sigma_{\mathbf{x}} \end{pmatrix}$  the covariance matrices of respectively  $W(\mathbf{x})$  and  $W(\mathbf{s}, \mathbf{x})$ . For the details of the calculations see Appendix A.

For  $\alpha = 1$ , we note that the expressions (15) and (16) are the ones obtained for the Schlather process (Dombry et al. (2013)). They are equivalent to those obtained by Ribatet (2013a).

Conditionally to  $Z(\mathbf{x}) = z$ , the Poisson point process  $\Phi^{(\alpha)}$  can be decomposed into two independent Poisson processes, such that  $\Phi^{(\alpha)} = \Phi_{\mathbf{x},z}^{(\alpha)+} \cup \Phi_{\mathbf{x},z}^{(\alpha)-}$ , with

$$\begin{aligned} \Phi_{\mathbf{x},z}^{(\alpha)-} &= \{\phi^{(\alpha)} \in \Phi^{(\alpha)} \mid \forall i = 1, \dots, k, \phi^{(\alpha)}(x_i) < z_i\} = \{\phi_j^{(\alpha)-}\}_{j \geq 1}, \\ \Phi_{\mathbf{x},z}^{(\alpha)+} &= \{\phi^{(\alpha)} \in \Phi^{(\alpha)} \mid \exists i = 1, \dots, k, \phi^{(\alpha)}(x_i) = z_i\} = \{\phi_j^{(\alpha)+}\}_{j \geq 1}. \end{aligned} \quad (17)$$

### 3.2. Algorithm

The algorithm proposed by Dombry et al. (2013) for the conditional simulation of Brown-Resnick and Schlather processes can easily be adapted for extremal  $t$  processes. It is based on the density  $\lambda_{\mathbf{x}}^{(\alpha)}$  and the decomposition of  $\Phi^{(\alpha)}$ . The following notations have been chosen to be consistent with notations of Dombry et al. (2013).

Let us note  $x_{\tau_j} = \{x_i \mid x_i \in \tau_j, i = 1, \dots, k\}$ ,  $z_{\tau_j} = \{z_i \mid Z(x_{\tau_j}) = z_i, i = 1, \dots, k\}$  and  $\tau_j^c$  the complementary of  $\tau_j$  in  $\{x_1, \dots, x_k\}$ .

**Step 1** Draw a random partition  $\tau$  of the set  $(x_1, \dots, x_k)$  conditionally to  $Z(\mathbf{x}) = z$ , according to the distribution:

$$\mathbb{P}[T = \tau \mid Z(\mathbf{x}) = z] \propto \prod_{j=1}^{|\tau|} \lambda_{\mathbf{x}_{\tau_j}}^{(\alpha)}(z_{\tau_j}) \int_{\{u_j < z_{\tau_j^c}\}} \lambda_{\mathbf{x}_{\tau_j^c} | \mathbf{x}_{\tau_j}, z_{\tau_j}}^{(\alpha)}(u_j) du_j. \quad (18)$$

**Step 2** Given a partition  $\tau = (\tau_1, \dots, \tau_l)$ , draw  $l$  independent extremal functions  $(\phi_1^{(\alpha)+}, \dots, \phi_l^{(\alpha)+})$  at  $\mathbf{s}$  with finite dimensional distribution

$$\mathbb{P}[\phi_j^{(\alpha)+}(\mathbf{s}) \in dv \mid Z(\mathbf{x}) = z, T = \tau] \propto \left\{ \int \mathbb{1}_{\{u_j < z_{\tau_j^c}\}} \lambda_{(\mathbf{s}, \mathbf{x}_{\tau_j^c}) | \mathbf{x}_{\tau_j}, z_{\tau_j}}^{(\alpha)}(v, u_j) du_j \right\} dv. \quad (19)$$

Set  $Z^+(s_j) = \max(\phi_1^{(\alpha)+}(s_j), \dots, \phi_l^{(\alpha)+}(s_j))$  for  $j \in 1, \dots, m$ .

**Step 3** Independently of the two previous steps, draw a number  $N$  sufficiently large of  $\phi^{(\alpha)-}$  at  $(\mathbf{s}, \mathbf{x})$  such that  $\phi^{(\alpha)-}(\mathbf{x}) \leq z$  and set

$$Z^-(s_j) = \max_{i \geq 1} \{\phi_i^{(\alpha)-}(s_j)\} \text{ for } j \in 1, \dots, m, \quad (20)$$

Then  $\tilde{Z}(\mathbf{s}) = \max\{Z^+(\mathbf{s}), Z^-(\mathbf{s})\}$  is distributed according to the conditional distribution of  $Z(\mathbf{s})$  given  $Z(\mathbf{x}) = z$ .

Step 1 is the most challenging one because of the tremendous number of partition candidates. Indeed the number of possible partitions is equal to the Bell number (Spivey (2008)). This number grows very fast with the number of conditioning points. For instance, for  $k = 10$ , it is already equal to 115 975. Direct calculation of the probability in equation (18) is therefore not possible for a realistic number of conditioning points. In order to bypass this computational issue, Dombry et al. (2013) propose drawing the partition using a Gibbs sampler (Casella and George (1992)) which changes the class partition of one observed value at a time. Then once the transition probabilities involved simplified, the chain runs through the space of all possible partitions and converges to the target law described in equation (18).

Nevertheless, there are still some issues that need to be solved. The major problem is that the integrals involved in the transition distribution (see equation (18)) have to



be calculated at each step of the Gibbs sampler. The dimension of these integrals can be as high as  $(k - 1)$ . The values of the integrals can be estimated by classic Monte-Carlo methods or quasi Monte-Carlo methods (Caffisch (1998)) but in both cases, the estimation requires a huge number of points to get satisfying results and the Gibbs sampling can become computationally heavy.

For Step 2, it is necessary to simulate from the distribution  $\lambda_{\mathbf{s}|\mathbf{x},z}^{(\alpha)}$ . Using the same arguments as in Dombry et al. (2013), it can be shown that  $\lambda_{\mathbf{s}|\mathbf{x},z}^{(\alpha)}(u)$ ,  $u \in \mathbb{R}^m$  is the density of a multivariate Student distribution with  $(k + \alpha)$  degrees of freedom, with location parameter

$$\mu = \Sigma_{\mathbf{s}:\mathbf{x}} \Sigma_{\mathbf{x}}^{-1} z, \quad (21)$$

and scale matrix

$$\tilde{\Sigma} = \frac{a_{\mathbf{x}}(z)}{k + \alpha} (\Sigma_{\mathbf{s}} - \Sigma_{\mathbf{s}:\mathbf{x}} \Sigma_{\mathbf{x}}^{-1} \Sigma_{\mathbf{x}:\mathbf{s}}) \text{ with } \Sigma_{\mathbf{s},\mathbf{x}} = \begin{pmatrix} \Sigma_{\mathbf{s}} & \Sigma_{\mathbf{s}:\mathbf{x}} \\ \Sigma_{\mathbf{x}:\mathbf{s}} & \Sigma_{\mathbf{x}} \end{pmatrix} \text{ and } a_{\mathbf{x}}(z) = z^T \Sigma_{\mathbf{x}}^{-1} z. \quad (22)$$

Therefore  $\mathbb{1}_{\{u_j < z_{\tau_j^c}\}} \lambda_{(\mathbf{s}, \mathbf{x}_{\tau_j^c}^c) | \mathbf{x}_{\tau_j}, z_{\tau_j}}^{(\alpha)}(\mathbf{v}, u_j)$  involved in the integral of (19) is a truncated multivariate Student distribution and step 2 results in a simulation under constraints. A procedure to perform this simulation is proposed and detailed in the following section.

Step 3 is the most accessible one. It consists in simulating values for an extremal  $t$  process at every conditioning and unconditioning location and check whether they respect the constraints given by the conditioning points using an acceptance/rejection procedure. In this context, the rejection rate of the algorithm is not very high even for a large number of conditioning points. Indeed, it rarely happens that one of the  $\phi_i^{(\alpha)-}$  becomes greater than the conditioning value, because in this model, the conditioning values are meant to be, by construction, extreme realizations of these functions.

To simulate  $Z^-$ , especially the Poisson point process  $\xi$ , we adapted the simulation techniques described in Schlather (2002) for the Schlather process by noting that

$$\xi = \left\{ \left( \sum_{i=1}^n e_i \right)^{-1/\alpha} \right\}_{n \in \mathbb{N}}, \quad (23)$$

with  $e_i \sim \mathcal{E}(1)$  a standard exponential distribution, is a Poisson point process with intensity measure  $\alpha t^{-(\alpha+1)} dt$ .

By construction the sequence  $\xi$  is decreasing and for  $i$  large enough  $\xi_i W_i(\cdot)$  will not participate to the maxima  $Z^-$  for any  $x \in \mathcal{X}$  provided that  $W_i(x) < C$  for all  $x \in \mathcal{X}$ .  $C$  can be set to 5, and we choose the number  $N$  of sub-extremal functions such that the probability that the last built one contributes to the maximum is lower than 0.001. This number depends on  $\alpha$  and increases with it.

### 3.3. Multivariate Student under constraints

A straightforward way to sample from a truncated multivariate Student distribution is to simulate from a multivariate Student distribution and add an acceptance/rejection procedure (Dombry et al. (2013)). Nevertheless, the rate of acceptance of this algorithm becomes very low for a number of conditioning points higher than 30. We propose using the general methodology developed by Geweke (1991) for sampling a  $p$ -dimensional vector  $y$  from a multivariate Student distribution with  $p$  linear constraints i.e.

$$y \sim t_\nu(\mu, \Sigma), \text{ with } a \leq Dy \leq b,$$

with  $D$  a full rank  $p \times p$  matrix and some elements of the  $p$ -dimensional vectors  $a$  and  $b$  can be equal respectively to  $-\infty$  and  $+\infty$ . This method involves two steps simulating alternatively the denominator and the numerator of the multivariate Student random vector. The denominator is a constrained  $\chi^2$  variable simulated via an acceptance/rejection procedure (on the ratio between the multivariate normal realization and the  $\chi^2$  variable) and the numerator is updated sequentially according to a Gibbs sampler. The whole procedure is repeated until convergence.

It has to be noted that the rejection rate of the  $\chi^2$  variable simulation is low because the constraint is univariate, and that the Gibbs sampler avoids using an acceptance/rejection procedure on vectors. The adaptation of this methodology to the step (2) (equation (19)) of the algorithm is straightforward setting for each  $j \in \{1, \dots, l\}$ , all  $a$  components to  $-\infty$ ,  $D$  to the  $(m + \#(\tau_j^c))$ -dimensional identity matrix, the first  $m$  components of  $b$  to  $+\infty$  and the  $\#(\tau_j^c)$  last ones to  $z_{\tau_j^c}$ . Consequently, all tools are available to perform conditional simulations of extremal  $t$  processes with this algorithm.

## 4. Comparison criteria

For assessing the quality of the results, criteria comparing distributions resulting from the conditional simulations and a value corresponding to an observation have to be used.

The accuracy of conditional simulations is assessed using first the continuous ranking probability score crps (Gneiting and Raftery (2007)). Letting  $F$  be the predictive cumulative distribution function derived from the  $n$  conditional simulations  $\{\hat{y}^{(j)}\}_{j=1, \dots, n}$  and  $y$  the test observation, the continuous ranking probability score is written

$$\text{crps}(F; y) = \int_{-\infty}^{+\infty} [F(t) - H(t - y)]^2 dt, \quad (24)$$

where  $H(u)$  is the Heavyside function which is equal to 0 for  $u < 0$  and 1 otherwise. The steeper the cdf is around the real value, the better. A good score is a score close to 0. It also appears that the calculation of the crps will result in a value with the same unit as the variable of interest.

Generally, there is no closed-form expression for this integral. Nevertheless, when  $F$  denotes a GEV distribution, a closed-form expression exists (Friederichs and Tho-

rarinsdottir (2012)).

Moreover, the conditional simulations are also evaluated based on the quantile score (qs) defined as

$$\text{qs}_p(F; y) = \nu_p(y - F^{-1}(p)) \quad (25)$$

where  $p$  is the probability of interest and  $\nu_p(u) = pu$  if  $u \geq 0$  and  $\nu_p(u) = (p - 1)u$  otherwise. As shown by Friederichs and Thorarinsdottir (2012), the crps is proportional to the integral of the  $\text{qs}_p$  over all possible probabilities  $p$ . Then a good  $\text{qs}_p$  is a  $\text{qs}_p$  close to 0 as well. By choosing a high probability  $p$ , the quantile score focuses on the predictive performance in the upper tail (Gneiting and Raftery (2007)).

Then, overall scores are built from equations (24) and (25), by averaging on all points:

$$\begin{aligned} \overline{\text{CRPS}}((F_1, \dots, F_m); (y_1, \dots, y_m)) &= \frac{1}{m} \sum_{i=1}^m \text{crps}(F_i; y_i). \\ \overline{\text{QS}}_p((F_1, \dots, F_m); (y_1, \dots, y_m)) &= \frac{1}{m} \sum_{i=1}^m \text{qs}_p(F_i; y_i). \end{aligned} \quad (26)$$

In addition to these two univariate scores, which assess the marginal qualities of the conditional simulation, we propose a multivariate score:

$$\text{MS}_d(\hat{y}_1, \dots, \hat{y}_n; y_1, \dots, y_m) = \sum_{I \in \mathcal{P}_m(d)} \sum_{j=1}^n \|\hat{y}_I^{(j)} - y_I\|_d \quad (27)$$

with  $\|\cdot\|_d$  the euclidean norm in  $\mathbb{R}^d$ ,  $\mathcal{P}_m(d)$  the ensemble of indices of length  $d$  in  $\{1, \dots, m\}$ ,  $y_I = \{y_i | i \in I\}$  and  $\hat{y}_I^{(j)} = \{\hat{y}_i^{(j)} | i \in I\}$ . This score enables us to assess the spatial coherency of the conditional simulations.

In order to compare models, skill scores can be built (Gneiting and Raftery (2007)):

$$S^{\text{skill}} = \frac{S^{\text{mod}} - S^{\text{ref}}}{S^{\text{opt}} - S^{\text{ref}}} \quad (28)$$

with  $S^{\text{mod}}$ ,  $S^{\text{ref}}$  and  $S^{\text{opt}}$  the score of the current model, the score of a reference model and the optimal score respectively. For models that improve on the reference model,  $S^{\text{skill}}$  belongs to  $[0, 1]$  with 1 correspond to the optimal case. For  $\overline{\text{CRPS}}$ ,  $\overline{\text{QS}}_p$  and  $\text{MS}_d$ ,  $\overline{\text{CRPS}}^{\text{opt}} = \overline{\text{QS}}_p^{\text{opt}} = \text{MS}_d^{\text{opt}} = 0$  and the corresponding skill scores can be written for these three scores as

$$\text{CRPSS} = \frac{\overline{\text{CRPS}}^{\text{ref}} - \overline{\text{CRPS}}^{\text{mod}}}{\overline{\text{CRPS}}^{\text{ref}}}, \text{QSS}_p = \frac{\overline{\text{QS}}_p^{\text{ref}} - \overline{\text{QS}}_p^{\text{mod}}}{\overline{\text{QS}}_p^{\text{ref}}}, \text{MSS}_d = \frac{\text{MS}_d^{\text{ref}} - \text{MS}_d^{\text{mod}}}{\text{MS}_d^{\text{ref}}}. \quad (29)$$

which corresponds to the percentages of  $\overline{\text{CRPS}}$ ,  $\overline{\text{QS}}_p$  or  $\text{MS}_d$  improvement brought by the current model with respect to the reference model.

## 5. Simulation Study

In order to check the properties of the conditional simulations provided by this algorithm, simulations according to extremal  $t$  processes are drawn, and conditional simulations at non conditioning points (test sample) are compared to their real values. 100 points are sampled at random in a unit square. 20 are kept as conditioning points, and the other 80 are used as test values. For all simulations the covariance function is the exponential with scale parameter 1.

### 5.1. Influence of the partition

As stated in the section 3, the first and hardest step for conditional simulation is to draw a partition of the conditioning points according to the distribution given in equation (18). Indeed, it appears that the Gibbs sampler converges very slowly to the target law and then it is likely that we do not draw the partition from the target distribution. We investigate how the resulting partition distribution, obtained at different convergence stages of the Gibbs sampler, influences the performance of the whole algorithm. For each  $\alpha \in \{2, 4, 6\}$ , we simulate extremal  $t$  fields and compare the performance of the algorithm when (i) the real partition is known, (ii) when it is drawn from a Gibbs sampler (10 000 iterations) according to the distribution (18) and (iii) when the partition is drawn randomly. Table 1 gives the CRPSS calculated by taking the random partition model as the reference.

$\alpha$	Real Part.			Gibbs sampler Part.			Random Part.		
	2	4	6	2	4	6	2	4	6
$\overline{\text{CRPS}}$	<b>0.158</b>	<b>0.132</b>	<b>0.118</b>	0.237	0.182	0.173	0.250	0.192	0.175
CRPSS	<b>36.8%</b>	<b>31.2%</b>	<b>32.8%</b>	5.1%	5.2%	1.2%	0	0	0

Table 1: Importance of a good estimation of the partition: values of the  $\overline{\text{CRPS}}$  and CRPSS for  $\alpha = 2, 4, 6$ , (i) when the real partition is known, (ii) when the partition is drawn from a Gibbs sampler (10 000 iterations) according to the distribution (18) and (iii) when the partition is drawn randomly.

When the partition is drawn from the Gibbs sampler, the CRPSS is slightly greater than 5% for  $\alpha=2$  and  $\alpha=4$  and around 1% for  $\alpha=6$ . However, it is greater than 30% when the real partition is used. This demonstrates the strong gain to be had by drawing close to the real one and confirms that the efforts made for this step of the algorithm are inevitable and should even be developed. Indeed, the gain in drawing the partition with the Gibbs sampler is not very high compared to using a random partition (less than 6%) and far from the 30% obtained with the real partition. We can think that drawing randomly the partition instead of performing a Gibbs sampler for time-saving reasons for instance, would not change considerably the results especially when the number of conditioning points is high.

### 5.2. Influence of $\alpha$

We investigate how the value of the parameter  $\alpha$  influences the quality of the conditional simulations. For each value of  $\alpha$ , 15 simulations of an extremal  $t$  process are

drawn and 20 conditioning points are sampled for each simulation. Then 50 conditional simulations are performed according to each conditioning point set. We suppose that the real partition is known, just as  $\alpha$  and  $\rho(\cdot)$  which are not estimated.

The conditional simulations of the extremal  $t$  process are compared to conditional simulations of a Schlather process (Schlather (2002)), a Brown-Resnick process (Brown and Resnick (1977)), and a Gaussian process with the same conditioning points and unconditional simulations of the extremal  $t$  process. Ordinary Kriging (Chilès and Delfiner (1999)) is set as the reference prediction for the CRPSS calculation. The conditional simulations of the Schlather and Brown-Resnick processes are performed according to the algorithm of Dombry et al. (2013) and with **SpatialExtremes** (Ribatet (2013b)). For the Gaussian process conditional simulations, we apply first a Gaussian margin transformation and then perform the methodology described in Lantuéjoul (2002) for Gaussian processes (even if it is only marginally Gaussian). For the ordinary kriging and the Gaussian simulations, the covariance function is estimated from the data.

Figure 1 describes the evolution of the different CRPSS with  $\alpha$ . As expected since it is based on the right model, the extremal  $t$  model gives better CRPSS than do the other models. It also appears that the unconditional extremal  $t$  simulations become more and more efficient as  $\alpha$  increases. Indeed, when  $\alpha$  increases, the variance of the extremal  $t$  process tends to 0 (Opitz (2013)) and the simulated samples show very little variability, making the information added by the conditional values less influential. It is worth noticing that Gaussian margin transformation followed by conditional simulations of Gaussian fields shows surprisingly good results and it appears to be very efficient specially when  $\alpha$  is high. The relevance of this approach in conditional simulation framework has also been stated in Oesting and Schlather (2014).

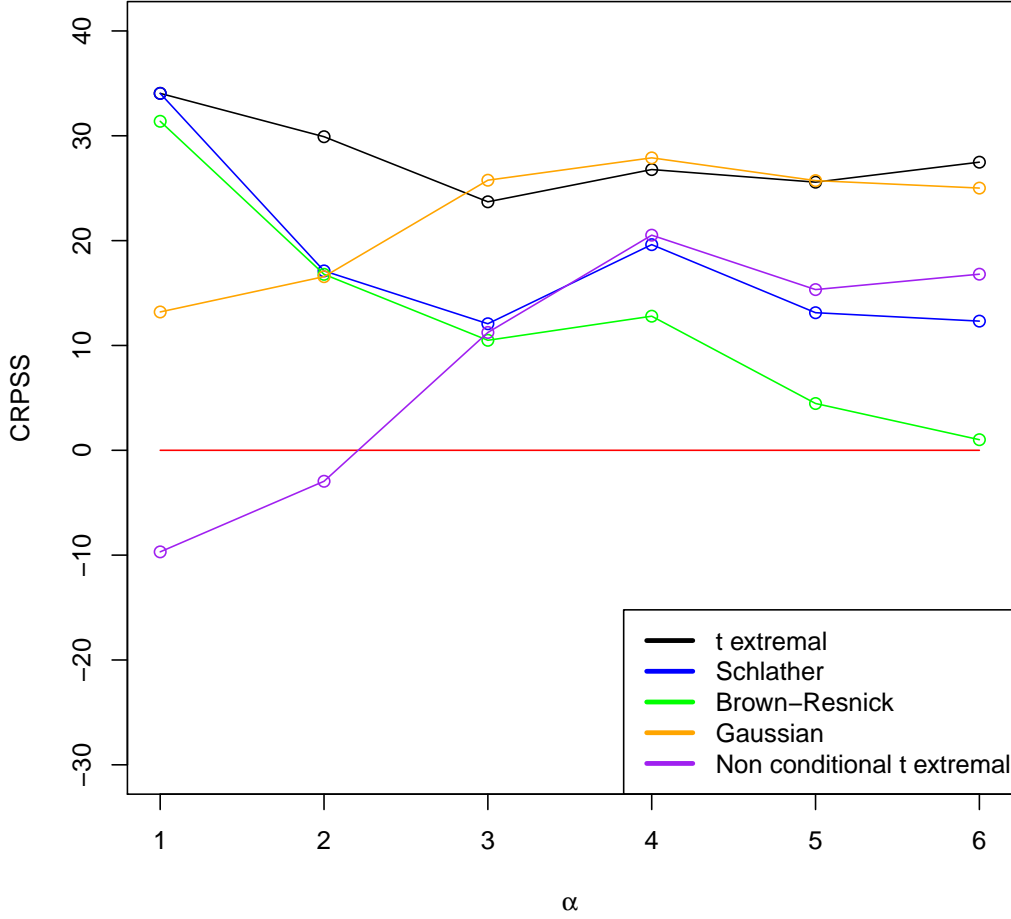


Figure 1: Evolution of the CRPSS (in %) with the  $\alpha$  with ordinary kriging as reference.

Nevertheless, for every values of  $\alpha$ , the CRPSS of the extremal  $t$  simulations is around 20% showing more stability in good performances than any other tested models.

### 5.3. Asymptotic dependence properties

We now investigate a notion related to the asymptotic dependence for the conditional simulations. As mentioned by Dombry et al. (2013),  $Z(\cdot)|\{Z(\mathbf{x}) = z\}$  is not max-stable and cannot be asymptotically dependent. However, the process integrated on all the conditional events recovers the max-stability property, and its asymptotic dependence can be characterized by the madogram (Cooley et al. (2006)). Using the madogram we compare the pairwise extremal coefficient estimates  $\theta_{est}$  to the theoretical extremal coefficient function  $\theta_{th}(h)$  at a distance  $h$  in order to check that the asymptotic dependence behavior of the conditional simulations is similar to that of the theoretical

one.

For each  $\alpha \in \{2, 4, 6\}$ , we simulate 1 000 extremal  $t$  fields at the same 20 conditional points and estimate the extremal coefficient function from the remaining 80 points with the R package `SpatialExtremes` (Ribatet (2013b)). Figure 2 compares the theoretical extremal function  $\theta_{th}(h)$  and the estimated one  $\theta_{est}(h)$  calculated from the madogram of the integrated conditional simulations.

For the three values of  $\alpha$ , the conditional simulations respect the expected asymptotic dependence behavior. Indeed, the theoretical extremal coefficient function fits relatively well the pairwise extremal coefficient estimates, although it tends to overestimate them when  $\alpha$  becomes high.

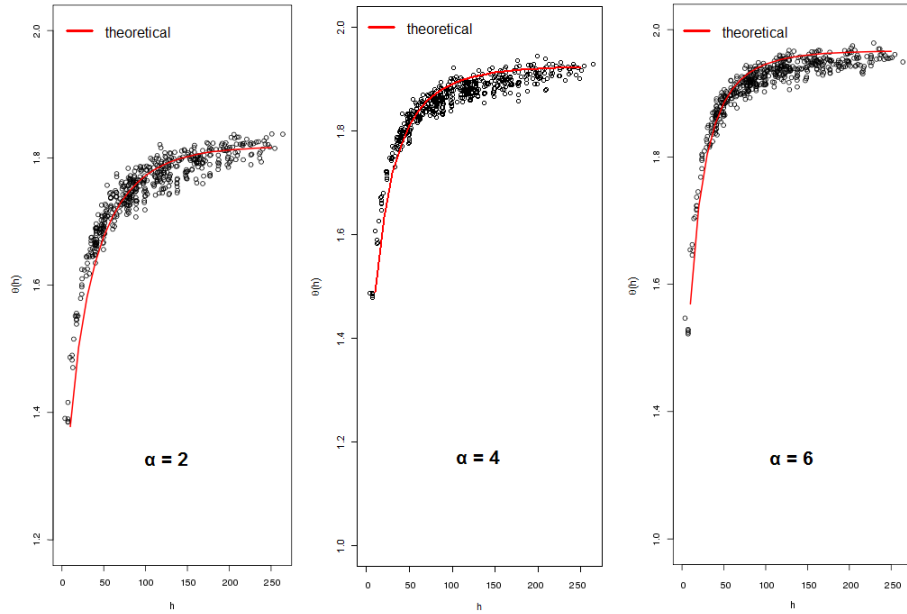


Figure 2: Extremal coefficient: comparison of  $\theta_{th}(h)$  and  $\theta_{est}(h)$  for  $\alpha \in \{2, 4, 6\}$ .

## 6. Application to precipitation data

### 6.1. Description and margin modeling

To check the behavior of the model and the algorithm on real data, we extracted daily precipitation data from the SAFRAN analyses (Quintana-Seguí et al. (2008)) in the southeast of France covering the region  $[3.4^\circ\text{E}, 6.0^\circ\text{E}] \times [43.2^\circ\text{N}, 44.5^\circ\text{N}]$  – as represented by the red square in Figure 3 – with a spatial resolution of  $8 \text{ km} \times 8 \text{ km}$ , which corresponds to 457 grid cells. Then, for each cell, we computed the autumnal maximum daily precipitation (from 91 days between September 1<sup>st</sup> and November 30<sup>th</sup>) for 1960 – 2007: hence, for each grid cells, we have 48 autumnal maximum daily precipitation values. 50 grid cells have been randomly selected as conditioning points. Conditional simulations are then performed on the 407 remaining grid cells given the

observation of the 50 conditioning spots.

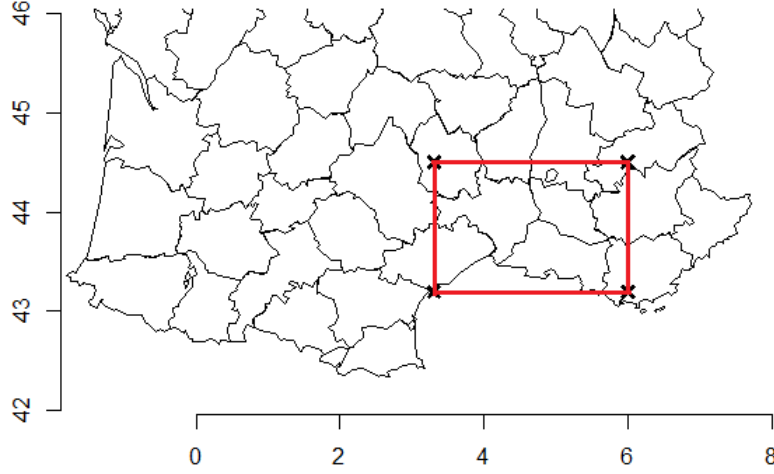


Figure 3: Study area of SAFRAN data subset.

The choice of the autumnal period allows us to make the temporal stationarity assumption, as it is commonly accepted that rainfall data can be considered as stationary within a given season.

In order to perform the conditional simulations presented in section 3 the margins have first to be transformed into an  $\alpha$ -Fréchet distribution. If  $Z$  is distributed according to a  $\text{GEV}(\mu, \sigma, \xi)$ , then  $Y$  defined as

$$Y = \left(1 + \frac{\xi(Z - \mu)}{\sigma}\right)_+^{1/\alpha\xi} \quad (30)$$

follows an  $\alpha$ -Fréchet distribution (Beirlant et al. (2004)).

To avoid the estimation of the parameters ( $\mu, \sigma$  and  $\xi$ ) of the GEV for all points and to deal with spatial non-stationarity, these parameters are modeled as functions of the spatial coordinates  $LAT$  and  $LON$  except  $\xi$  that is assumed to be constant for the whole region:

$$\begin{aligned} \mu &= \exp[\mu_0 + \mu_1 * LAT + \mu_2 * LON] \\ \sigma &= \exp[\sigma_0 + \sigma_1 * LAT + \sigma_2 * LON] \\ \xi &= \xi_0 \end{aligned} \quad (31)$$

It can be observed that the exponential formulation is employed for  $\mu$  and  $\sigma$  as  $\sigma$  has to be positive for model definition reasons and  $\mu$  is chosen to be positive because of the nature of the observed phenomenon (precipitation).



### 6.2. Inference

The parameters of the model, i.e.:

1.  $\mu_0, \mu_1, \mu_2, \sigma_0, \sigma_1, \sigma_2, \xi$ , the parameters for the margin model,
2.  $\alpha$ , the extremal  $t$  parameter,
3. and  $a$  the scale parameter of the exponential correlation function of  $W$ ,

have to be estimated from the data. Nevertheless, for dimensions higher than 2, the multivariate density of the extremal  $t$  process is not tractable and usual likelihood maximization methods are not available. Alternatively, composite likelihood (Lindsay (1988), Varin et al. (2011)) can be used. It consists in calculating bivariate densities for each pair of points and maximizing their weighted sum. Let us note  $\{z_{m,x_i}\}$  the autumnal maximum on  $m = 1, \dots, M$  years for a given  $x_i$  in  $(x_1, \dots, x_k)$ . The composite log-likelihood is

$$\mathcal{LC}(\Psi; \mathbf{z}) = \sum_{m=1}^M \sum_{i=1}^{k-1} \sum_{j=i+1}^k w_{ij} l_{ij}(\Psi) \quad (32)$$

with  $l_{ij}(\Psi) = \log f(z_{m,x_i}, z_{m,x_j}; \Psi)$  the bivariate log-density for points  $x_i$  and  $x_j$  with  $\Psi$  the vector of parameters to estimate, and  $w_{ij}$  the corresponding weights. These weights  $w_{ij}$  have been chosen to be decreasing (exponentially) with the distance between  $x_i$  and  $x_j$ . The maximization of this composite likelihood leads to consistent estimations of the target parameters (Padoan et al. (2010), Xu and Reid (2011)).

### 6.3. Results

The inference has been driven by the data over the whole time period. The maximization of the composite likelihood gives  $\alpha = 2.2$  (with  $\theta_\infty = 1.83$ ) which means that even for long distances there is still some asymptotic dependence. Furthermore, the estimated range of the exponential covariance function is equal to 360 km. The estimated parameters and the confidence intervals are presented in Table B.1.

The validation is performed only for the year 2007 with 100 conditional simulations. The most represented partition in the history of the Gibbs sampler is composed of 5 classes. However, there is no notable spatialization of these classes (not shown).

Figure 4 shows the individual crps scores (as defined in equation (24)) computed for the year 2007 taking the observed values as test observations ( $y_i$ ) and the empirical distributions of the conditional simulations as the predictive cdf ( $F_i$ ).

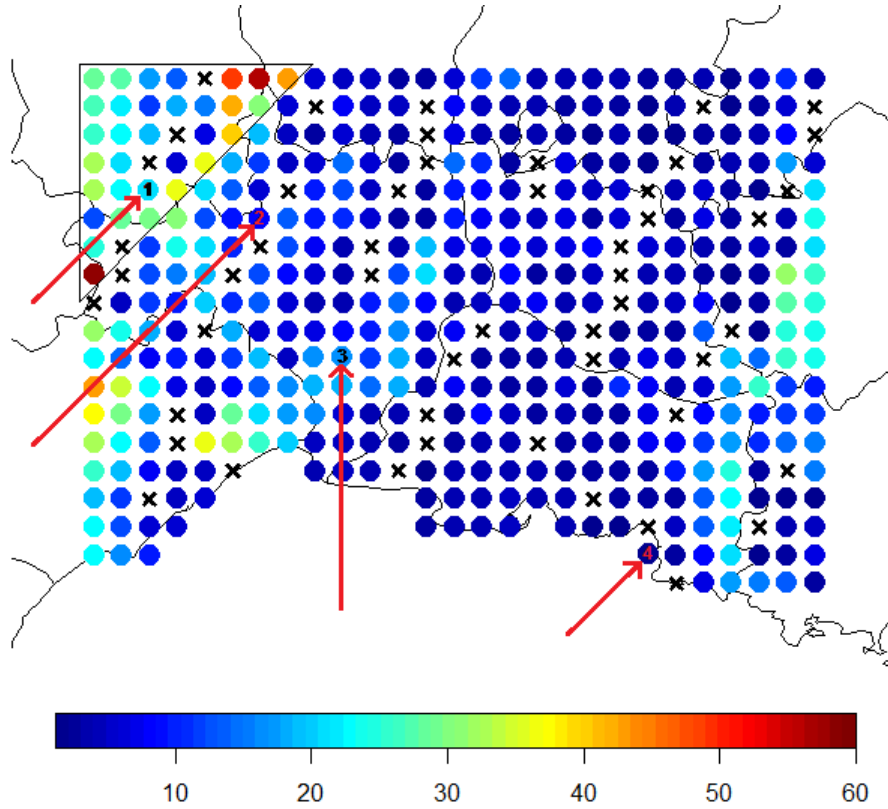


Figure 4: Map of the crps (in mm) for 2007. For further illustrations, four grid cells have been arbitrarily selected and are indicated here (from 1 to 4).

In addition, Figure 5 describes the distributions of the conditional simulations at four arbitrarily chosen grid cells specified on Figure 4. It appears naturally that the closer to conditioning points the new location is, the better the simulation (e.g. points 2 and 4 are close to conditioning points whereas point 3 is far). However, this is clearly not the only important criterion. Indeed, point 1 is quite close to conditioning points yet its simulations are not as good as expected.

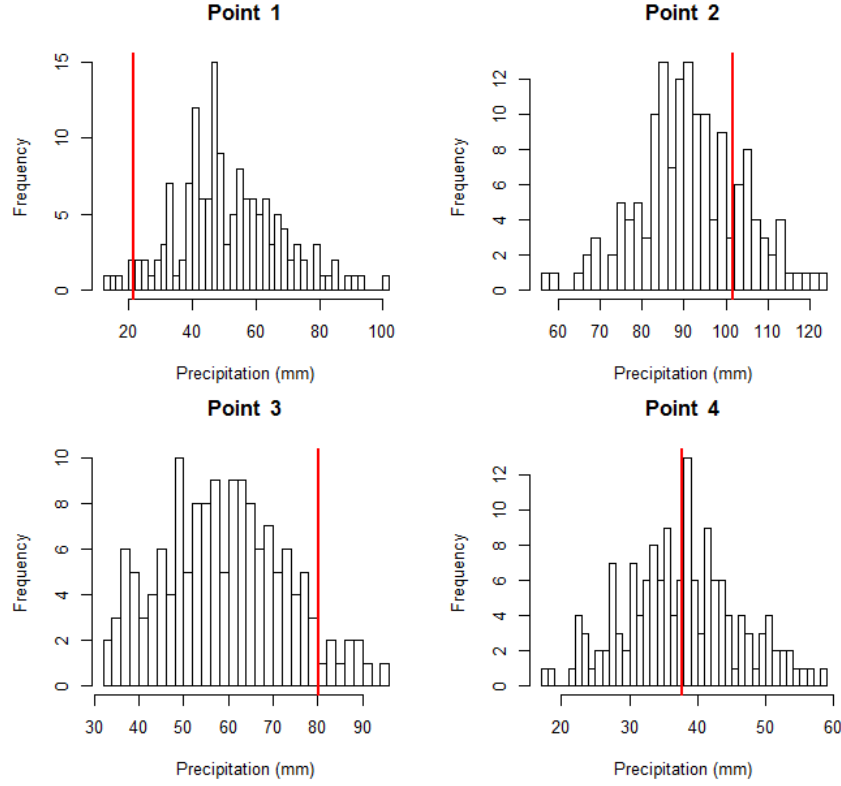


Figure 5: Histograms of the conditional simulations at 4 locations, with the real value (in red).

In order to evaluate the quality of these conditional simulations, other models and algorithms are considered and compared:

- (A) conditional simulation of extremal  $t$  model, with Gibbs sampler,
- (B) conditional simulation of extremal  $t$ , with random partitions,
- (C) analog model,
- (D) historical model,
- (E) unconditional simulation of extremal  $t$  model.

For simulations based on the analog model (C) (e.g. Barnett and Preisendorfer (1978); Zorita and Von Storch (1999) or Pascal et al. (2013)), at each non conditioning location we consider the nearest conditioning grid cell. Then we look at the history of this point (1960-2006) to find the year with the closest observed value: the value observed this year at the non conditioning location is taken as the estimate value for the non conditioning point. For the so-called “historical” model (D), we randomly draw a value in the history of a non conditioning location and consider it as the value of the simulation.

Table 2 gives the  $\overline{\text{CRPS}}$ ,  $\text{CRPSS}$ ,  $\overline{\text{QS}}_{95}$ ,  $\text{QSS}_{95}$ ,  $\text{MS}_2$  and  $\text{MSS}_2$  for each method with the historical model (D) as reference.

Method	$\overline{\text{CRPS}}$	$\text{CRPSS}$	$\overline{\text{QS}}_{95}$	$\text{QSS}_{95}$	$\text{MS}_2$	$\text{MSS}_2$
(A) Cond. Sim. Gibbs	10.05	-4.1%	<b>1.95</b>	<b>34.8%</b>	27.55	31.3%
(B) Cond. Sim. Random	10.96	-13.1%	2.02	32.4%	28.99	27.7%
(C) Analog	10.53	-9.1%	5.68	-90.0%	<b>18.41</b>	<b>54.1%</b>
(D) Historical	<b>9.65</b>	<b>0%</b>	2.99	0%	40.11	0%
(E) Uncond. Sim.	12.71	-31.6%	3.33	-11.4%	48.07	-19.9%

Table 2:  $\overline{\text{CRPS}}$ ,  $\text{CRPSS}$ ,  $\overline{\text{QS}}_{95}$ ,  $\text{QSS}_{95}$ ,  $\text{MS}_2$  and  $\text{MSS}_2$  for the methods (A), (B), (C), (D) (as reference) and (E).

The historical model appears as the best one regarding the  $\overline{\text{CRPS}}$  score. In term of  $\text{MS}_2$ , the model analog, which respects by construction the spatial coherency, gives the best results. Model (A) shows good results in terms of  $\overline{\text{QS}}_{95}$  and generally appears to be the most versatile method adapted to different purposes (for each score, it gets at worst the second best value). It is also important to note that method (A) is better than method (B): there is a real gain in running a Gibbs sampler to estimate the partitions.

Generally, models (A)-(B) are the best to catch high quantile behavior (see  $\overline{\text{QS}}_{95}$ ). Nevertheless, as crps is the overall score of  $qs_p$  integrated over all probabilities  $p$ , it indicates that for lower probabilities, model (A) is not as good as the others. It appears that for  $p \geq 0.75$ ,  $\overline{\text{QS}}_p$  of model (A) calculated over the whole region is the best.

Furthermore, by deepening the analysis of the results, it appears that for some points the error (in terms of crps) is fairly high, especially for model (A). Those points are located in the upper left corner of Figure 4 (i.e., the northwest of the region), corresponding to the higher region(s) of the Cevennes mountains. The climate characteristics in this area are very different from the rest of the region and it is likely that the chosen covariates (LAT and LON) are not sufficient to overcome the inhomogeneity of this region.

Then, for comparison purposes inside the stationary framework, we removed these points for the calculation of the scores. Table 3 shows the new scores.

Method	$\overline{\text{CRPS}}$	$\text{CRPSS}$	$\overline{\text{QS}}_{95}$	$\text{QSS}_{95}$	$\text{MS}_2$	$\text{MSS}_2$
(A) Cond. Sim. Gibbs	<b>8.85</b>	<b>4.4%</b>	<b>1.78</b>	<b>37.2%</b>	30.14	21.8%
(B) Cond. Sim. Random	9.46	-2.1%	1.80	36.5%	30.79	20.0%
(C) Analog	10.14	-9.4%	5.55	-95.7%	<b>18.15</b>	<b>52.9%</b>
(D) Historical	9.27	0%	2.83	0%	38.51	0%
(E) Uncond. Sim.	11.13	-20.3%	3.09	-9.1%	47.24	-22.7%

Table 3:  $\overline{\text{CRPS}}$ ,  $\text{CRPSS}$ ,  $\overline{\text{QS}}_{95}$ ,  $\text{QSS}_{95}$ ,  $\text{MS}_2$  and  $\text{MSS}_2$  for the methods (A), (B), (C), (D)(as reference) and (E) with removal of the outliers area.

The CRPSS is now better for method (A), with a gain of 4.4% with respect to the reference model and both  $QSS_{95}$  and  $MSS_2$  remain good. It also appears that the  $\overline{CRPS}$  and CRPSS have been improved for all methods.

It is also interesting to check whether the conditional simulations follow the temporal variations of the real values. To this aim, we have simulated 50 conditional fields on 50 grid cells from the SAFRAN database, on the period 2000 – 2007 with the same estimated parameters. As running the Gibbs sampler for each year with 50 conditional points would be too time-consuming, the partitions are drawn randomly for each year. Figure 6 shows the temporal evolutions of the means of the conditional simulations at the four previously chosen grid cells.

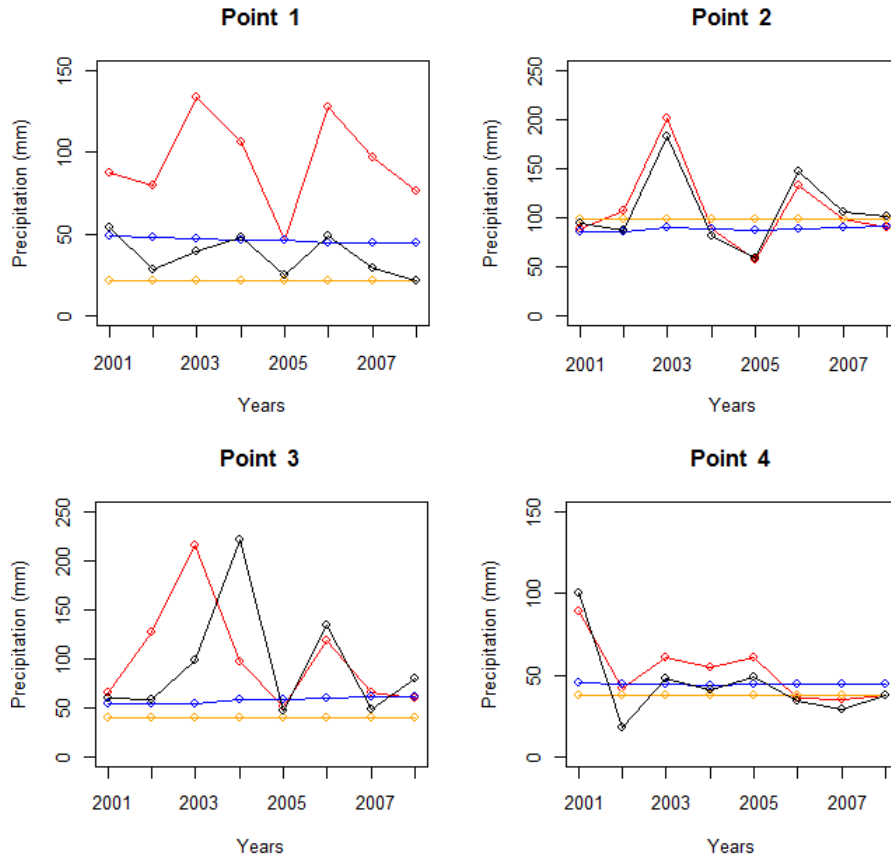


Figure 6: Comparison of the evolution of the mean of the conditional simulations (in red), the historical model (in blue), analog model (in orange) and the test values (in black).

For points 2 and 4, the temporal evolutions of the conditional simulations are almost the same as for the test values, whereas for point 3, there is a big discrepancy in the first period and a good match in the second period. A constant significant bias is noted for point 1, which is a location in the already mentioned mountain region, and it can be suspected that it behaves in a quite different way than do the conditioning points

even though it approximately follows the temporal variations.

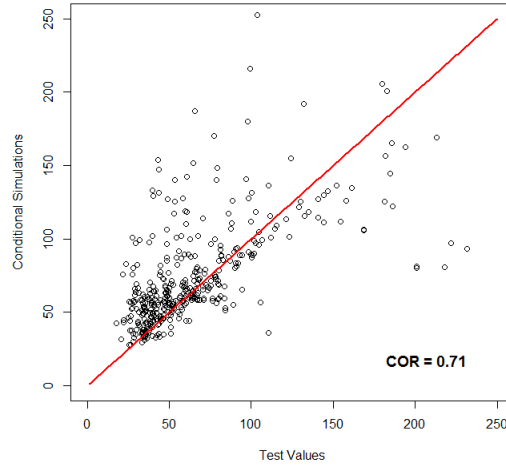


Figure 7: Correlation between the simulated values and the test values.

The scatterplot of the simulated values vs. the test values is represented in Figure 7. The Pearson correlation coefficient is equal to 0.71. It appears that the conditional simulations roughly follow the intrinsic variations of the data over the years. Indeed, a correlation coefficient of 0.71 is a satisfying result which indicates that the method is able to reproduce the general evolution of the observed phenomenon. In comparison, the correlation coefficient is 0.47 for the historical model and 0.33 for the analog model.

## 7. Discussion and conclusions

We have investigated conditional simulations for maxima precipitation using the extremal  $t$  model. Thanks to its parametrization, this model offers a great flexibility in modeling the asymptotic dependence. The conditional simulation algorithm designed by Dombry et al. (2013) for Brown-Resnick and Schlather models has been adapted to the extremal  $t$  process and the second step, consisting in simulating from a multivariate Student distribution under constraints, has been improved by a two-state Gibbs sampler proposed by Geweke (1991) enabling us to handle a greater number of conditioning points.

A simulation study highlighted that the methodology of conditional simulations for the extremal  $t$  process is effective for a large range of  $\alpha$  ( $\in [1, 6]$ ). Moreover, when the value of  $\alpha$  is too high, its estimation becomes more and more uncertain. In such cases, we would advise the use of the conditional simulations of Gaussian processes with transformed Gaussian margins (Oesting and Schlather (2014), Lantuéjoul (2002)) which are faster than the general Brown-Resnick processes (Kabluchko et al. (2009)) which show similar properties to those of an extremal  $t$  process for high values  $\alpha$ .

A major aspect that needs to be improved is the way to generate the partitions of

the conditioning points. Indeed, the Gibbs sampler requires too many computational resources to provide better results. Some alternative methods have to be considered in order to avoid the calculations of the multiple integrals involved in the transition probabilities.

The application performed on a dataset of rainfall in the south of France showed that the model is relevant and that the algorithm performs well, provided that the spatial and temporal assumptions of stationarity are fulfilled. The margins modelling and transformation (equations (33) and (34)) intended to make the transformed process spatially stationary. However, the sole use of the coordinates (LAT+LON) is likely not sufficient to remove all the spatial heterogeneity of the process especially at the boundaries or where some other major climatic factors have to be taken into account. For this example, a small sub-region with different climate characteristics has been identified where the spatial coordinates do not seem to be sufficient to get satisfying results. Hence, the introduction of other covariates, such as altitude or information from climate or meteorological situations, could help to improve the simulations.

However, for simulations of fields of maxima precipitation in a future climate change context, no conditioning points (i.e., really observed at locations) are available. The only information stems from Global Climate Models (GCM) or Regional Climate Models (RCM) driven by various greenhouse gas emission scenarios. Those models provide many data but at low spatial resolution (from about 300 km down to about 10 km at best, so far). Hence, the best those models can give will be mean precipitation values over relatively large regions. Those data are clearly not adapted to feeding hydrological, ecological or, more generally, environmental impact models, which require high-resolution input climate data. Hence, the use of a max-stable process to simulate data for the future and provide those simulations to environmental impact models needs to be based on a change of spatial resolution. In other words, the max-stable process will not have “station” values available as conditioning points but only large-scale climate conditioning information. Therefore, one important need is to be able to condition max-stable processes by those large-scale climate data, which corresponds to performing a “statistical downscaling” to generate high-resolution fields of extreme precipitation. This downscaling in a spatial modeling framework will enable us to statistically generate fields and times series that will be useful for many impact studies trying to anticipate the potential consequences of extreme precipitation events in a climate change context.

## Appendix A. Calculation of the density and the conditional density

*Proof.* According to equation (13), we have

$$\begin{aligned}\Lambda_{\mathbf{x}}^{(\alpha)}(A) &= \int_0^{+\infty} \mathbb{P}(m_{\alpha}^{-1/\alpha} \xi W(\mathbf{x}) \in A) \xi^{-(\alpha+1)} d\xi \\ &= \int_0^{+\infty} \int_{\mathbb{R}^k} \mathbb{1}_{\{m_{\alpha}^{-1/\alpha} \xi z \in A\}} g_{\mathbf{x}}(z) dz \xi^{-(\alpha+1)} d\xi\end{aligned}\tag{A.1}$$

with for  $z \in \mathbb{R}^k$ ,

$$g_{\mathbf{x}}(z) = (2\pi)^{-(k/2)} |\Sigma_{\mathbf{x}}|^{-1/2} \exp\left[-\frac{1}{2} z^T \Sigma_{\mathbf{x}}^{-1} z\right].$$

Setting  $y = m_{\alpha}^{-1/\alpha} \xi z$ , we get

$$\Lambda_{\mathbf{x}}^{(\alpha)}(A) = \int_0^{+\infty} \int_A g_{\mathbf{x}}\left(\frac{y}{\xi m_{\alpha}^{-1/\alpha}}\right) \frac{1}{m_{\alpha}^{-k/\alpha}} \xi^{-(k+1+\alpha)} dy d\xi \quad (\text{A.2})$$

and then

$$\begin{aligned} \lambda_{\mathbf{x}}^{(\alpha)}(y) &= \int_0^{+\infty} g_{\mathbf{x}}\left(\frac{y}{\xi m_{\alpha}^{-1/\alpha}}\right) m_{\alpha}^{k/\alpha} \xi^{-(k+\alpha+1)} d\xi. \\ &= (\sqrt{2\pi})^{-k} |\Sigma_{\mathbf{x}}|^{-1/2} m_{\alpha}^{k/\alpha} \int_0^{+\infty} \exp\left[-\frac{1}{2m_{\alpha}^{-2/\alpha} \xi^2} y^T \Sigma_{\mathbf{x}}^{-1} y\right] \xi^{-(k+\alpha+1)} d\xi \end{aligned} \quad (\text{A.3})$$

By noting that for  $l \geq 2$ ,

$$\int_0^{+\infty} \exp\left[-\frac{z^T \Sigma_{\mathbf{x}}^{-1} z}{2m_{\alpha}^{-2/\alpha} \xi^2}\right] \xi^{-l} d\xi = 2^{(l-3)/2} [z^T \Sigma_{\mathbf{x}}^{-1} z]^{-(\frac{l-1}{2})} m_{\alpha}^{-(l-1)/\alpha} \Gamma\left(\frac{l-1}{2}\right) \quad (\text{A.4})$$

$\lambda_{\mathbf{x}}^{(\alpha)}(\mathbf{z})$  can be written:

$$\begin{aligned} \lambda_{\mathbf{x}}^{(\alpha)}(z) &= \pi^{-\frac{k-1}{2}} |\Sigma_{\mathbf{x}}|^{-1/2} [z^T \Sigma_{\mathbf{x}}^{-1} z]^{-(\frac{k+\alpha}{2})} \Gamma\left(\frac{k+\alpha}{2}\right) \Gamma((\alpha+1)/2)^{-1} \\ &= \frac{\pi^{\frac{\alpha-1}{2}}}{w_{k+\alpha-2}} |\Sigma_{\mathbf{x}}|^{-1/2} [z^T \Sigma_{\mathbf{x}}^{-1} z]^{-(\frac{k+\alpha}{2})} \Gamma((\alpha+1)/2)^{-1} \end{aligned} \quad (\text{A.5})$$

and

$$\lambda_{\mathbf{s}|\mathbf{x},z}^{(\alpha)}(u) = \frac{\lambda_{\mathbf{s},\mathbf{x}}^{(\alpha)}(u, z)}{\lambda_{\mathbf{x}}^{(\alpha)}(z)} = \frac{w_{k+\alpha-2}}{w_{k+m+\alpha-2}} \sqrt{|\Sigma_{\mathbf{s}:\mathbf{x}} \Sigma_{\mathbf{x}}^{-1}|} \frac{[z^T \Sigma_{\mathbf{x}}^{-1} z]^{(k+\alpha)/2}}{[(u, z)^T \Sigma_{\mathbf{s},\mathbf{x}}^{-1} (u, z)]^{(k+m+\alpha)/2}} \quad (\text{A.6})$$

with  $w_d = \pi^{d/2} / \Gamma(d/2 + 1)$  the volume of the unit ball in  $d$  dimensions.  $\square$



## Appendix B. Estimated parameters

Parameters	Estimation	CI (95%)
$\mu_0$	-3.0565	[ -3.0942 ; -3.0187 ]
$\mu_1$	-0.2227	[ -0.2233 ; -0.2221 ]
$\mu_2$	0.1804	[ 0.1803 ; 0.1805 ]
$\sigma_0$	-4.3431	[ -4.3732 ; -4.3129 ]
$\sigma_1$	-0.3665	[ -0.3675 ; -0.3654 ]
$\sigma_2$	0.2084	[ 0.2084 ; 0.2085 ]
$\xi$	0.1145	[ 0.1128 ; 0.1163 ]
$\alpha$	2.2077	[ 2.0354 ; 2.3945 ]
$a$	1.2262	[ 1.1304 ; 1.3301 ]

Table B.1: Estimated parameters (and CI) with composite likelihood maximization.

## References

- Barnett, T., Preisendorfer, R., 1978. Multifield analog prediction of short-term climate fluctuations using a climate state vector. *Journal of Atmospheric Sciences* 35, 1771–1787.
- Beaulant, A.-L., Joly, B., Nuissier, O., Somot, S., Ducrocq, V., Joly, A., Sevault, F., Deque, M., Ricard, D., 2011. Statistico-dynamical downscaling for mediterranean heavy precipitation. *Quarterly Journal of the Royal Meteorological Society* 137 (656), 736–748.
- Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J., Waal, D. D., Ferro, C., 2004. *Statistics of extremes. Theory and applications*. Hoboken, NJ: John Wiley & Sons.
- Blanchet, J., Davison, A. C., 09 2011. Spatial modeling of extreme snow depth. *The Annals of Applied Statistics* 5 (3), 1699–1725.
- Brown, B. M., Resnick, S. I., 1977. Extreme values of independent stochastic processes. *Journal of Applied Probability* 14 (4), 732–739.
- Buishand, T., de Haan, L., Zhou, C., 2008. On spatial extremes: with application to a rainfall problem. *The Annals of Applied Statistics*, 624–642.
- Cafisch, R. E., 1 1998. Monte carlo and quasi-monte carlo methods. *Acta Numerica* 7, 1–49.
- Casella, G., George, E. I., 1992. Explaining the gibbs sampler. *The American Statistician* 46 (3), 167–174.

- Chilès, J.-P., Delfiner, P., 1999. *Geostatistics: Modeling Spatial Uncertainty*. Wiley, New-York.
- Coles, S., Tawn, J., 1994. Statistical methods for multivariate extremes - an application to structural design. *Journal of the Royal Statistical Society - Series C: Applied Statistics* 43 (1), 1–48.
- Cooley, D., Naveau, P., Poncet, P., 2006. Variograms for spatial max-stable random fields. In: *Dependence in probability and statistics*. Springer, pp. 373–390.
- Davison, A. C., Gholamrezaee, M. M., 2012. Geostatistics of extremes. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 468 (2138), 581–608.
- Davison, A. C., Padoan, S. A., Ribatet, M., 05 2012. Statistical modeling of spatial extremes. *Statistical Science* 27 (2), 161–186.
- De Haan, L., 1984. A spectral representation for max-stable processes. *The Annals of Probability* 12 (4), 1194–1204.
- De Haan, L., De Ronde, J., 1998. Sea and wind: multivariate extremes at work. *Extremes* 1 (1), 7–45.
- Dombry, C., Eyi-Minko, F., 2013. Regular conditional distributions of continuous max-infinitely divisible random fields. *Electron. J. Probab* 18 (7), 1–21.
- Dombry, C., Éyi-Minko, F., Ribatet, M., 2013. Conditional simulation of max-stable processes. *Biometrika* 100 (1), 111–124.
- Ducrocq, V., Nuissier, O., Ricard, D., Lebeaupin, C., Thouvenin, T., 2008. A numerical study of three catastrophic precipitating events over southern france. ii: Mesoscale triggering and stationarity factors. *Quarterly Journal of the Royal Meteorological Society* 134 (630), 131–145.
- Fawcett, L., Walshaw, D., 2012. Estimating return levels from serially dependent extremes. *Environmetrics* 23 (3), 272–283.
- Friederichs, P., Thorarinsdottir, T. L., 2012. Forecast verification for extreme value distributions with an application to probabilistic peak wind prediction. *Environmetrics* 23 (7), 579–594.
- Fuentes, M., Henry, J., Reich, B., 2013. Nonparametric spatial models for extremes: application to extreme temperature data. *Extremes* 16 (1), 75–101.
- Gaume, J., Eckert, N., Chambon, G., Naaim, M., Bel, L., 2013. Mapping extreme snowfalls in the french alps using max-stable processes. *Water Resources Research* 49 (2), 1079–1098.

- Geweke, J., 1991. Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints and the evaluation of constraint probabilities. In: Computing science and statistics: Proceedings of the 23rd symposium on the interface. Citeseer, pp. 571–578.
- Gneiting, T., Raftery, A. E., 2007. Strictly proper scoring rules, prediction, and estimation.
- IPCC, 2011. Managing the risks of extreme events and disasters to advance climate change adaptation. Tech. rep., Intergovernmental Panel of Climate Change (IPCC).
- Kabluchko, Z., Schlather, M., de Haan, L., 09 2009. Stationary max-stable fields associated to negative definite functions. *The Annals of Probability* 37 (5), 2042–2065.
- Lantuéjoul, C., 2002. Geostatistical simulation: models and algorithms. Springer.
- Lebeaupin, C., Ducrocq, V., Giordani, H., 2006. Sensitivity of torrential rain events to the sea surface temperature based on high-resolution numerical forecasts. *Journal of Geophysical Research: Atmospheres* 111 (D12), n/a–n/a.
- Lindsay, B. G., 1988. Composite likelihood methods. *Contemporary Mathematics* 80 (1), 221–39.
- Mariethoz, G., Renard, P., Cornaton, F., Jaquet, O., 2009. Truncated plurigaussian simulations to characterize aquifer heterogeneity. *Ground Water* 47 (1), 13–24.
- Oesting, M., Schlather, M., 2014. Conditional sampling for max-stable processes with a mixed moving maxima representation. *Extremes* 17 (1), 157–192.
- Opitz, T., 2013. Extremal t-processes: Elliptical domain of attraction and a spectral representation. *Journal of Multivariate Analysis* 122, 409 – 413.
- Padoan, S. A., 2011. Multivariate extreme models based on underlying skew-t and skew-normal distributions. *Journal of Multivariate Analysis* 102, 977–991.
- Padoan, S. A., Ribatet, M., Sisson, S. A., 2010. Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association* 105 (489), 263–277.
- Pascal, Y., Tamara, S., Philippe, D., Laurent, M., Robert, V., Mathieu, V., 2013. Ensemble reconstruction of the atmospheric column from surface pressure using analogues. *Climate dynamics* 41 (5-6), 1333–1344, eng.
- Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., Morel, S., 2008. Analysis of near-surface atmospheric variables: Validation of the safran analysis over france. *Journal of applied meteorology and climatology* 47 (1), 92–107.

- Ribatet, M., 2013a. Spatial extremes: Max-stable processes at work. *Journal de la Societe Francaise de Statistique* 154 (2), 156–177.
- Ribatet, M., 2013b. *Spatialextremes : Modelling spatial extremes*.
- Schlather, M., 2002. Models for stationary max-stable random fields. *Extremes* 5 (1), 33–44.
- Schlather, M., Tawn, J. A., 2003. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* 90 (1), 139–156.
- Smith, R. L., 1990. Max-stable processes and spatial extreme, unpublished.
- Spivey, M. Z., 2008. A generalized recurrence for bell numbers. *Journal of Integer Sequences* 11 (2), Article ID 08.2.5, 3 p., electronic only–Article ID 08.2.5, 3 p., electronic only.
- Varin, C., Reid, N. M., Firth, D. D., 2011. An overview of composite likelihood methods. *Statistica Sinica* Vol.21 (No.1), 5–42.
- Vrac, M., Drobinski, P., Merlo, A., Herrmann, M., Lavaysse, C., Li, L., Somot, S., 2012. Dynamical and statistical downscaling of the french mediterranean climate: uncertainty assessment. *Natural Hazards and Earth System Science* 12 (9), 2769–2784.
- Vrac, M., Yiou, P., 2010. Weather regimes designed for local precipitation modeling: Application to the mediterranean basin. *Journal of Geophysical Research: Atmospheres* 115 (D12), 2156–2202.
- Wang, Y., Stoev, S. A., 2011. Conditional sampling for spectrally discrete max-stable random fields. *Advances in Applied Probability* 43 (2), 461–483.
- Xu, X., Reid, N., 2011. On the robustness of maximum composite likelihood estimate. *Journal of Statistical Planning and Inference* 141 (9), 3047–3054.
- Zorita, E., Von Storch, H., 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *Journal of climate* 12 (8), 2474–2489.